

Role of Hidden Neurons in an Elman Recurrent Neural Network in Classification of Cavitation Signals

Ramadevi R

Sathyabama University,
Jeppiaar Nagar, Rajiv Gandhi Road, Chennai 600119, India

Sheela Rani B

Sathyabama University,
Jeppiaar Nagar, Rajiv Gandhi Road, Chennai 600 119, India

Prakash V

Indira Gandhi Centre for Atomic
Research, Kalpakkam,
Chennai 603 102, India

ABSTRACT

This paper is intended to present the outcome of a study conducted on the cavitation data collected from accelerometer which is installed at the down stream of the cavitation test loop, to illustrate that the hidden neurons in an ANN modelling tool, indeed, do have roles to play in percentage of classification of cavitation signal. It sheds light on the role of the hidden neurons in an Elman Recurrent type ANN model which is used to classify the cavitation signals. The results confirmed that the hidden-output connection weights become small as the number of hidden neurons becomes large and also that the trade-off in the learning stability between input-hidden and hidden-output connections exists. The Elman recurrent network propagates data from later processing stage to earlier stage. A copy of the previous values of the hidden units are maintained which allows the network to perform sequence-prediction. In the present work, the optimum number of hidden neurons is evolved through an elaborate trial and error procedure. It is concluded that our approach has a significant improvement in learning and also in classification of cavitation signals.

General Terms

Optimum ANN model development.

Keywords

ANN, Elman Recurrent Network, Hidden neurons, Activation Function, Learning Algorithm, Cavitation.

1. INTRODUCTION

Artificial Neural networks have been successfully applied to classification problems. Multilayer networks with back propagation learning algorithm is limited to searching for a appropriate network topology including number of hidden neurons and weights to solve the learning problems in hand. Too small networks are unable to adequately learn the problem while excessively large networks tend to over fit the training data and consequently results in poor performance [1]. A major problem lies in specifying the size of the neural network. Even for moderately size networks the number of parameters may become large compared to the number of data [2]. Most practical learning problems are known to be computationally complex and hard to optimize. In this study, network performance is examined as a function of number of hidden neurons, because the number of hidden neurons in a multi layered neural network is subject to debate. The accuracy of an

ANN model varies with the number of hidden neurons. Often in an ANN application the number of hidden neurons is selected based on the particular application and then only after testing various numbers of hidden neurons. In our investigation, we observed the influence of the number of hidden neurons by comparing the relative performance of ANNs with increasing the hidden neurons in various patterns. This is carried out using a seven layered Elman recurrent network with resilient back propagation algorithm. The paper is organized as follows, Section II describe about signal acquisition and data sets analyzed, section III describes ANN modeling module for classification of various cavitation stages, section IV analyses the performance results and section V concludes with future work.

2. SIGNAL ACQUISITION AND DATA SETS ANALYZED

PFBR is a sodium cooled pool type reactor. PFBR core subassemblies are supported vertically inside the sleeves provided in the Grid plate. The Grid Plate acts as a coolant header through which flow is distributed among the Subassembly to remove fission heat. Since the power profile of the reactor core is not uniform, it is necessary to distribute the coolant flow to each subassembly according to their power levels. PFBR core is divided into 15 zones (Zone I to Zone XV) such as fuel, blanket, reflector, storage etc according to their respective power levels. To achieve the maximum temperature of sodium at core outlet, it is essential to feed the subassemblies with a flow proportional to the power generation. The flow zoning in the different Subassembly's of the reactor core is achieved by installing pressure dropping devices in the foot of the subassembly. Orifices having Honey-comb type geometry were developed to meet the flow zoning requirements of Fuel zone. Hydraulics of these orifices is important in achieving the required pressure drop without cavitation [3]. Cavitation free performance of these devices has to be ensured because cavitation leads to reactivity perturbations, positive sodium void coefficient, dimensional changes due to erosion etc. Cavitation test has been carried out to find the various cavitation stages. Vibration data has been collected from accelerometers which are placed downstream side of orifices of all zones for two different flow rates viz 110% and 100% [4].

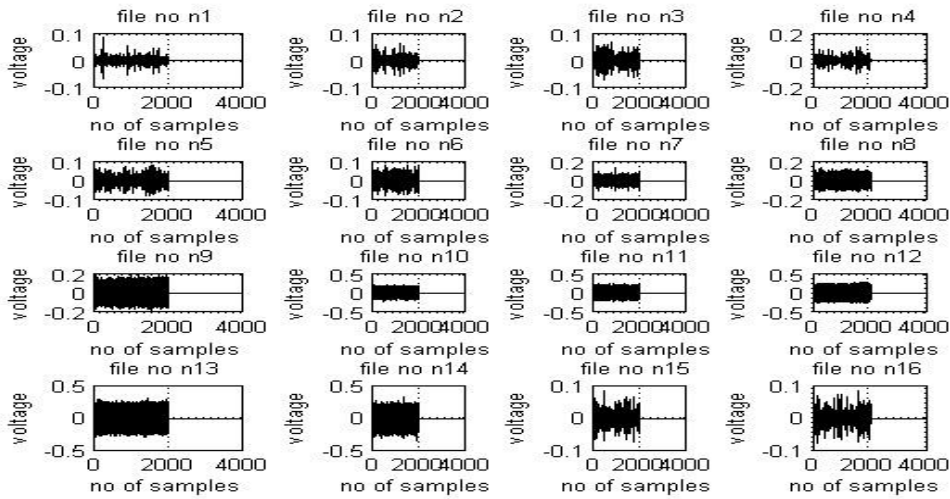


Figure1. Few Example Signals from Zone II

Four data sets have been analyzed viz zone II, Zone IV, Zone VI and Zone VII with channel 1 & 2. Where Channel 1 means 110% flow rate and channel 2 means 100% flow rate.

Each signal containing 2002 samples. Input data are provided in table 1. Figure 1 shows few example data sets from Zone II

Table1. Input Data

Zone	Channel	No. of Signals
II	1	58
IV	1	78
	2	78
VI	1	28
	2	15
VII	1	68
	2	68

3. ANN MODELLING MODULE

3.1 Network

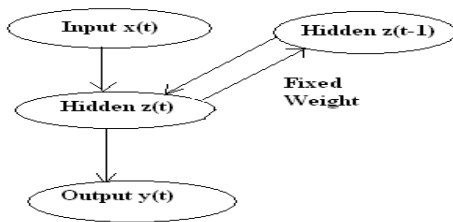


Figure2. Architecture of Recurrent Network

A recurrent neural network (RNN) is a class of neural network where connections between units form a directed cycle. This creates an internal state of the network. While a feed forward network propagates data linearly from input to output, recurrent networks propagate data from later processing stages to earlier stages.

Figure 2 shows architecture of an Elman recurrent network, with the addition of a set of context units in the input layer. There are connections from the hidden layer to these context units fixed

with a weight. At each time step, the input is propagated in a standard feed-forward fashion, and then a learning rule is applied. The fixed back connections result in the context units always maintaining a copy of the previous values of the hidden units (since they propagate over the connections before the learning rule is applied). Thus the network can maintain a sort of state, allowing it to perform the tasks as sequence-prediction.

In this approach a seven layered Elman recurrent network is trained through resilient back propagation algorithm. Elman network needs more hidden neurons in its hidden layer than are actually required for a solution by another method. While a solution may be available with fewer neurons, the Elman network is less able to find the most appropriate weights for hidden neurons since the error gradient is approximated. Therefore, having a fair number of neurons to begin with makes it more likely that the hidden neurons will start out dividing up the input space in useful ways. Elman networks can be trained with either of two functions, train or adapt. When using the function train to train an Elman network at each epoch, the entire input sequence is presented to the network, and its outputs are calculated and compared with the target sequence to generate an error sequence. For each time step, the error is back propagated to find gradients of errors for each weight and bias. This gradient is actually an approximation since the contributions of weights and biases to errors via the delayed recurrent connection are ignored. This gradient is then used to update the weights with the back prop training function chosen by the user [5].

3.2 Algorithm

Resilient back Propagation algorithm is a local adaptive learning scheme performing supervised batch learning in feed forward neural networks. The basic principle of this algorithm is to eliminate the harmful influence of the size of the partial derivative on the weight step. RPROP modifies the size of the weight step taken adaptively, and the mechanism for adaptation in RPROP does not take into account the magnitude of the gradient as seen by a particular weight, but only the sign of the gradient (positive or negative). This allows the step size to be

adapted without having the size of the gradient interfere with the adaptation process [1]. Resilient Back propagation Algorithm is generally much faster than the standard steepest descent algorithm. It also has a very good feature that it requires only a modest increase in memory requirements. It is a systematic method to train the neural network. The purpose of it is to eliminate the harmful effects of the magnitudes of the partial derivatives. Only the sign of the derivative is used to determine the direction of the weight update and the magnitude of the derivative has no effect on the weight update [6].

4. PERFORMANCE ANALYSIS

The magnitude of accelerometer signal increases with the development of cavitation. So, magnitude has been chosen as a feature inputs to neural network. The classification range has been obtained through vigorous analysis of various cavitation signals. The classification range for No cavitation is - 0.009 to 0.09, incipient cavitation is 0.1 to 0.99 and Developed cavitation is 1 to 2.9.

In the present work the optimum network architecture evolved out through an elaborate trial and error procedure. The network parameters such as Network structures, number of layers, and number of neurons in each layer, transfer function of layers, learning function and performance evaluation has been varied to find the best network parameters.

The training function can be any of the back propagation training functions such as trainlm, trainbfg, trainrp, traingd, etc. Here, trainrp has been selected as training function.

The learning function learnqdm has been chosen for this application to dynamically treat all weights in similar manner and to avoid a situation of no learning. So that the network will not end in global minima which will adversely affect the system performance and will result in massive errors and misclassifications.

Too small learning rate will cause the system to take a very long time to converge and too large will cause to oscillate indefinitely and diverge [7]. So, learning rate has been chosen as 0.01.

For training data sets totally 15 signals, 5 signal on each category (5 no cavitation signals, 5 Incipient cavitation signals and 5 developed cavitation signals) has been chosen from various zones as train data sets. The remaining signals are used as test data sets. For detecting various stages of cavitation, at first newelm structure has been studied with resilient back propagation algorithm. The effects of number of hidden neurons have been studied. The results are provided in the table 2.

Number of neuron in each layer is important and if it is low, neural network cannot reflect nonlinear mapping between input and output. On the other hand, if they are more than required, the network produces nonlinear mapping with unsatisfactory performance. Number of hidden layer neurons of seven layer network has been studied by trial and error analysis method. It is clear that seven layer recurrent network with 50, 40, 30, 20, 15 and 10 as hidden neurons has high percentage of cavitation detection.

Table2. Effect of Hidden Neurons

S.No	Network	% of Cavitation detection
1	[12, 10, 8, 6, 4, 2, 1]	38.4
2	[30, 25, 20, 15, 10, 5, 1]	45.8
3	[30, 20, 15, 10, 5, 2, 1]	56.3
4	[40, 30, 20, 15, 10, 5, 1]	53.2
5	[50, 40, 30, 20, 15, 10, 1]	61.0
6	[60, 50, 40, 30, 25, 20, 1]	59.1
7	[100, 80, 60, 40, 20, 10, 1]	48.7
8	[160, 80, 40, 20, 10, 5, 1]	40.1
9	[320, 160, 80, 40, 20, 10, 1]	55.7

Table 2 shows the effect of varying number of hidden neurons on the classification of cavitation signals.

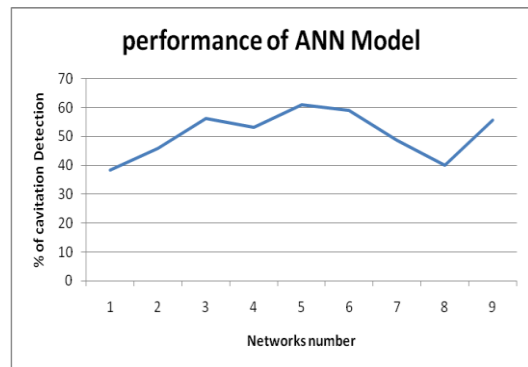


Figure3. Influence of Hidden Neurons on Performance of ANN model

Fig.3 presents the results from the investigation into the influence of the number of hidden neurons on performance of ANN model. It shows the performance of the ANN model in terms of classification rate. The trend of the plot indicates that the performance of the ANN with [50, 40, 30, 20, 15, 10, 1] hidden neuron is the best. A sharp decrease in performance is observed for the [100, 80, 60, 40, 20, 10, 1] & [160, 80, 40, 20, 10, 5, 1] ANNs. The parameter excitation function also has been considered in our research when hidden neurons are examined. Symmetric functions are chosen as excitation functions because these functions will cause the system to learn faster. Mostly nonlinear excitation function is introduced to maximize efficiency of multilayer networks [8, 9]. In this paper, combination of various excitation functions such as logsig, tansig, purelin have been studied. Results have been tabulated in table 3. Results suggest that the best excitation function for input layer is tansig, for all hidden layers it is logsig and purelin for output layer.

Therefore by analysis, the proposed neural network suggested for classification of various cavitation stages is,

```
Net = newelm(minmax(p), [50,40,30,20,15,10,1],
{'tansig','logsig','logsig','logsig','logsig','logsig','purelin'},
'trainrp', 'learnqdm', 'mse');
```

Table3. Effect of Excitation function with MSE Error function for Proposed Model

S.No	Excitation Function composition	Mean Square Error (%)
1	Tansig–Tansig–Tansig–Tansig–Tansig–Tansig–Purelin	-0.0687
2	Tansig–Logsig–Logsig–Logsig–Logsig–Logsig–Purelin	-0.0008
3	Logsig–Tansig–Tansig–Tansig–Tansig–Tansig–Purelin	0.7378
4	Logsig–Logsig–Logsig–Logsig–Logsig–Logsig–Purelin	0.1981
5	Tansig–Tansig–Purelin–Logsig–Logsig–Tansig–Purelin	0.0936

Testing of network has been carried out zone wise. Untrained inputs have been tested by proposed network and their efficiency was determined for MSE error function. The excitation functions used for the layers are tansig (for input), logsig (for all hidden layers) and purelin (for output). The optimum number of neurons used in each hidden layer is 50, 40, 30, 20, 15 and 10 respectively. trainrp was used with Learning rate = 0.01; Momentum constant = 0.9; Minimum performance gradient = 1e-10 as training algorithm. Common goal (0.471601) was fixed irrespective of zones and network was trained with 15 signals which are chosen from all zones and contains various types of cavitation and the rest of the signals were given for testing.

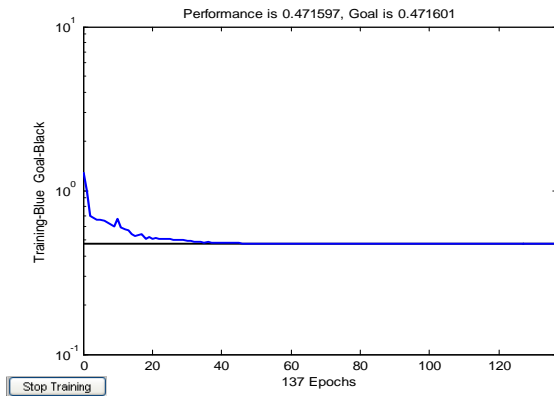


Figure4. Performance goal of Zone VI Channel 1 D57 (No cavitation)

Table4. Performance Goal with MSE error function

EPOCH	GOAL	MSE
0	0.471601	2.573619
25		0.592486
50		0.564801
75		0.540245
100		0.509762
125		0.467099
137	Goal Met	0.471597

Figure 3 shows the performance goal of training phase of zone VI channel 1 no cavitation signal (D57). The corresponding values are tabulated in Table 4. The network was trained and tested and the following results were obtained.

Table 5 shows performance analysis, with amplitude as input and trained using resilient back propagation algorithm using data from all zones of various cavitation types. The efficiency of the network has been tested on all zones. The analysis is carried out irrespective of zones (generalized).

5. CONCLUSION AND FUTURE WORK

The lower numbers of hidden neurons were expected to perform the best because networks with lower numbers of hidden neurons usually have the advantage of generalized predictions, but this was not the case. Results are presented on the classification accuracy. Significant improvement in classification rate is achieved using 50, 40, 30, 20, 15, 10 hidden neurons for seven layered Elman network model. Deciding how many hidden neurons to use is a complex task which this research tried to clarify. The above analysis concludes that the combination of seven layers with 50, 40, 30, 20, 15, 10 as number of hidden neurons and the combination of activation function Tansig (input), Logsig (hidden layers), Purelin (output) with Mean Squared Error (MSE) as Performance Function has been determined as the best Elman recurrent network parameters for classifying the various stages of cavitation signals. The overall percentage of cavitation detection for train data is 64.93% and for test data was found to be 65.34%. The Percentage of Detection PoD can be improved by proper selection of training signals from various zones and network parameters.

Future work can focus on role of activation function of the same network by fixing remaining other network parameter same.

Table5. Performance Analysis
Train Data Percentage of Detection: 64.93%
(Totally 15 signals, 5 signals on each category from all zones)

ZONE	CHANNEL	% OF DETECTION
		Test Data
II	1	61.21
	2	59.13
IV	1	56.72
	2	59.13
VI	1	68.97
	2	100
VII	1	56.44
	2	54.85
Over all %		65.34%

6. REFERENCES

- [1] Nidal F. Shilbayeh and Mahmoud Z. Iskandarani “Effect of Hidden Layer Neurons on the Classification of Optical Character Recognition Typed Arabic Numerals”. Journal of Computer Science 4 (7):578 – 584, 2008.

- [2] Parekh, R., J. Yang and V. Honavar, “Constructive neural network learning algorithms for multi-category pattern classification”, IEEE Trans. Neural Networks, 11: 436-451, 2000.
- [3] G.K. Pandey, D.Ramadasu, P.Anup Kumar, M.Thirumalai, G. Padmakumar, V. Prakash, C. Anandbabu, P.Kalyanasundaram, G.Vaidyanathan “Hydraulics of Honeycomb Type Orifices for Flow Zoning in FBR” Recent Advances in Networking, VLSI and Signal processing, pp 225 – 229.
- [4] P.K. Gupta, P.A. Kumar, A. Kaul, G.K. Pandey, G. Padmakumar, V. Prakash and C. Anandbabu “Neural Network Based Methodology for Cavitation Detection in Pressure Dropping Devices of PFBR” Proc. of national Seminar on Non destructive Evaluation Dec. 7 - 9, 2006, Hyderabad.
- [5] R. Rajesh, S. Chattopadhyay, M. Kundu “Prediction of Equilibrium Solubility of Co₂ in Aqueous Alkanolamines through Artificial Neural Network” Chemeca’06 17 – 20 September 2006.
- [6] Neural Network Toolbox, Signal Processing Toolbox, Statistics Toolbox of MATLAB.
- [7] K. Kumarci, M. Abdollahian, F. Kumarc, P.K Dehkordy, “Calculation of Natural Frequency of Arch Shape using Neural Network” CSCE 2008 Annual Conference, SCGC June 10 – 13, 2008.
- [8] Kazuyuki hara, kenji nakayama “Effects of activation functions in multi layer neural network for noisy pattern classification”, Kanazawa University, Japan.
- [9] Steven J.Burian, S.Rocky Durrans, Stephan J.Nix, Robert E.Pitt “Training Artificial Neural Networks to Perform Rainfall Disaggregation” Journal of Hydrologic Engineering, Vol.6, No.1, pp 43 – 51, 2001.